NEW ZEALAND STATISTICAL ASSOCIATION

The New Zealand Statistical Association was incorporated in 1949 with the object of "the encouragement of theoretical and applied statistics in New Zealand." In 1992 this object was expanded as a mission to "lead New Zealand to value and make intelligent use of statistical thinking and good statistical practice." Many of the early members used statistics as a tool in another discipline, there being no formal professional training offered anywhere in the country at that time. Today membership is still open to all, and in 1997 there are about 400 members. Formal accreditation of members has been discussed, but rejected in favor of a voluntary Code of Ethics. Organizations interested in the objects of the Association may join as corporate members. Initially the annual conference was the Association's only activity. Papers were often of low technical level, but they provided an opportunity for the few professionals to share their expertise in the ensuing discussions. The annual conference remains a focus for the Association's activities today, and although the technical level of the papers is now most respectable, the tradition of reaching out to nonstatisticians continues by holding joint conferences.

Education has been and still is a major concern of the Association. It has promoted statistics in schools and technical training, and has been and continues to be involved in designing curricula. It has published teaching materials and computer software, and each year awards statistics prizes at 23 secondary-school science fairs held throughout the country.

Concern over misuse of statistics in public affairs led to the establishment of a Survey Appraisal and Public Questions Committee in 1981. This committee is free to respond to approaches from groups suffering from what they judge to be a misuse of statistics. As a notable example, a 1995 report showed that family welfare benefit levels were based on a quite inadequate subset of the national household income and expenditure survey. The publicity given to the Committee's impartial, critical comments has improved the practice of survey sampling* in New Zealand generally.

The *New Zealand Statistician* first appeared in 1966 as a newsletter for the Association but it included some technical articles. In 1985 the *Newsletter* was separated from the *New Zealand Statistician,* which became a journal, publishing principally technical articles describing applications of statistics in New Zealand. In 1998 this journal amalgamated with the *Australian Journal of Statistics**. The combined *Australian and New Zealand Journal of Statistics* is published by Blackwells, and the first edition appeared in 1998. The combined journal has a New Zealand Associate Editor responsible for an applied statistics section, which is appropriate given the historical emphasis of the Association.

The Association is run by an Executive Committee of 16 who are scattered around the country, but it operates mainly by endorsing and supporting those members who have the interest to promote statistics in a special way. This has led to a varied range of initiatives over the years, an example of which was the publication of a book celebrating the role of women in statistics in New Zealand's Womens Suffrage Centennial Year, 1993. Regular subcommittees through the last decade have been those for education, survey appraisal and public questions, publications, science fairs, young statisticians, and history.

Through affiliation with the International Statistical Institute*, the Association maintains close relations with a number of statistical societies around the world, particularly the Statistical Society of Australia*. It is also a member body of the Royal Society of New Zealand (RSNZ). As such, the Association has a representative on the RSNZ Mathematical and Information Sciences Standing Committee, which provides it with a voice on national science policy.

JEFFREY J. HUNTER

NICKED-SQUARE DISTRIBUTION

The nicked-square (NS) distribution provides a valuable tool for studying the distribution theory of certain measures of agreement*. The

NS distribution is defined in this section, and the motivation for its development is discussed in the next.

Define the following sets of points in $\Re^2$: $S_0 = [0,1] \times [0,1]$, $S_1 = [0.4, 0.45] \times [0, 0.5]$, $S_2 = [0.55, 0.6] \times [0, 0.5]$, and $S_3 = [0.4, 0.6] \times [0, 0.5]$. The NS distribution, shown in Fig. 1, has the density

$$f(x,y) = \begin{cases} 1 & \text{if } (x,y) \in S_0 \backslash S_3, \\ 2 & \text{if } (x,y) \in (S_1 \cup S_2), \\ 0 & \text{otherwise.} \end{cases}$$

Here the random variables X and Y are dependent but uncorrelated. Apart from the nicked area, the NS distribution resembles the uniform square (US) distribution with density 1 on S_0 and 0 elsewhere.

MOTIVATION

Researchers may choose to study the agreement between two bivariate continuous measurements by constructing 2×2 contingency tables* (*see* TWO-BY-TWO TABLES) with categories defined by the two empirical medians of the marginal data (*see* EMPIRICAL BIVARIATE QUANTILE-PARTITIONED (EBQP) DISTRIBUTION.) Blomqvist [1] derived an asymptotic variance for such 2×2 tables* under certain regularity conditions, but subsequent work by Borkowf et al. [2] revealed that his asymptotic theory was correct only in special cases. They developed the NS distribution as a counterexample to illustrate differences between Blomqvist's asymptotic theory and their corrected asymptotic theory.

The EBQP distribution describes the distribution of $r \times c$ tables with categories defined by empirical quantiles* [2]; for 2×2 tables the extended hypergeometric (XH) distribution [4] has the same asymptotic distribution as that given by Blomqvist [1].

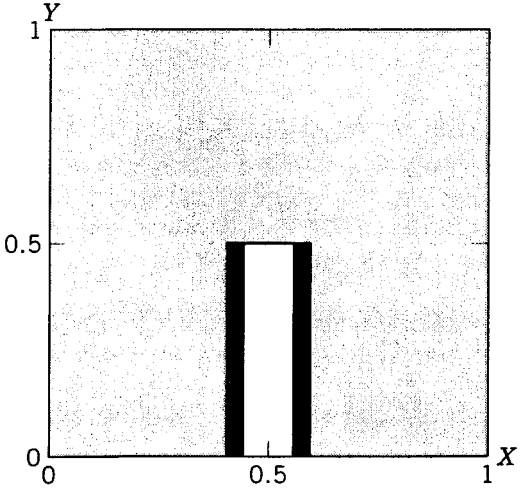

Figure 1 Nicked-square distribution. The density equals 1 in the gray region, 2 in the black region, and 0 elsewhere.

2 × 2 TABLES PARTITIONED BY EMPIRICAL MEDIANS

Suppose one samples N observations ($N = 2n$) from a bivariate continuous distribution $F(x,y)$ with marginal distributions $G(x)$ and $H(y)$ and conditional distributions $G(x \mid y)$ and $H(y \mid x)$. Let U and V be the empirical medians of the X and Y variables, respectively. Then the (X, Y) data can be partitioned into a 2×2 contingency table defined by these empirical medians with cell counts $\{m_{ij}\}$ $(i, j = 1, 2)$, as in Table 1; m_{11} denotes the number of observations with X and Y values that fall below both empirical medians.

Because this table has categories defined by the empirical medians, it has the EBQP distribution and its counts satisfy the constraints

$$m_{11} = m_{22} = n - m_{12} = n - m_{21}. \quad (1)$$

Thus, the table has a single degree of freedom.

Let $p_{ij} = m_{ij}/N$ and $\pi_{ij} = \lim_{N \to \infty} p_{ij}$. Let $\xi = G^{-1}(\frac{1}{2})$ and $\psi = H^{-1}(\frac{1}{2})$ denote the population medians of X and Y, respectively. Then $\pi_{11} = F(\xi, \psi)$ and $E[p_{ij}] \to \pi_{ij}$. Let $\gamma = G(\xi \mid \psi)$ and $\eta = H(\psi \mid \xi)$ denote the conditional proportions. From the asymptotic normal theory for $N^{1/2}(p_{11} - \pi_{11})$ and the

Table 1 2 × 2 Table of Counts Partitioned by Empirical Medians

	$Y < V$	$Y \geq V$	Total
$X < U$	m_{11}	m_{12}	n
$X \geq U$	m_{21}	m_{22}	n
Total	n	n	N

marginal constraints in (1), one can derive the asymptotic normal distributions of the cell counts $\{m_{ij}\}$, the empirical proportions $\{p_{ij}\}$, and measures of agreement calculated from 2×2 tables. The delta method* can be used to calculate the variances of such measures of agreement.

Since 2×2 EBQP tables have only one degree of freedom, all measures of agreement that are linear combinations of the cell counts of these tables are equivalent. For example, Cohen's kappa*, $\hat{\kappa} = 2[(p_{11} + p_{22}) - \frac{1}{2}]$, by the marginal constraints in (1), reduces to $\hat{\kappa} = 4p_{11} - 1$.

THREE ASYMPTOTIC VARIANCES

For 2×2 EBQP tables, Blomqvist [1] derived an asymptotic variance that equals that for 2×2 XH tables,

$$\mathrm{Var}_{\mathrm{XH}}(N^{1/2}p_{11}) \to \pi_{11}\left(\tfrac{1}{2} - \pi_{11}\right), \quad (2)$$

which differs from the asymptotic variance for 2×2 multinomial (MULT) tables,

$$\mathrm{Var}_{\mathrm{MULT}}(N^{1/2}p_{11}) \to \pi_{11}(1 - \pi_{11}). \quad (3)$$

One constructs 2×2 MULT tables similar to Table 1 but with random marginal totals by partitioning the original data by the population medians instead of by the empirical medians. In turn, one obtains 2×2 XH tables by selecting only those MULT tables that satisfy the marginal constraints in (1).

In the case of 2×2 EBQP tables partitioned by empirical medians,

$$\mathrm{Var}_{\mathrm{EBQP}}(N^{1/2}p_{11}) \to \pi_{11}(1 - \pi_{11}) + \tfrac{1}{4}(\gamma - \eta)^2 + \pi_{11}(2\gamma\eta - \gamma - \eta), \quad (4)$$

which differs from (2) and (3) in general. In the special case where the conditional proportions γ and η satisfy $\gamma = \eta = \frac{1}{2}$, this formula reduces to (2). Many common bivariate continuous distributions, such as the bivariate normal (regardless of correlation) and the US distribution, meet this condition; it always holds when X and Y are independent.

For example, the US distribution on S_0 has $\xi = \psi = \frac{1}{2}$, $\pi_{11} = \frac{1}{4}$, and $\gamma = \eta = \frac{1}{2}$. By contrast, the NS distribution, which was designed to violate this condition, has $\xi = \psi = \frac{1}{2}$ and $\pi_{11} = \frac{1}{4}$, but $\gamma = \frac{1}{2}$ and $\eta = 0$.

RESULTS

For the NS distribution, the asymptotic variances of $N^{1/2}p_{11}$ in 2×2 EBQP, XH, and MULT tables are $\frac{1}{8}$, $\frac{1}{16}$, and $\frac{3}{16}$, respectively [2]. This demonstrates that Blomqvist's variance, which equals that for XH tables, differs from the correct variance for EBQP tables. Similarly, the corresponding asymptotic variances for $N^{1/2}\hat{\kappa}$ are 2, 1, and 1, showing that the variance formula for $\hat{\kappa}$ given by Fleiss et al. [3], though correct for MULT tables, is incorrect for EBQP tables in general.

By contrast, for the US distribution, the asymptotic variances of $N^{1/2}p_{11}$ in 2×2 EBQP, XH, and MULT tables are $\frac{1}{16}$, $\frac{1}{16}$, and $\frac{3}{16}$, respectively, which reflects the equivalence of EBQP and XH tables when $\gamma = \eta = \frac{1}{2}$. The corresponding asymptotic variances for $N^{1/2}\hat{\kappa}$ are all 1.

Acknowledgment

This entry was prepared while C.B.B. held a National Research Council-National Institutes of Health Research Associateship at the National Heart, Lung, and Blood Institute's Office of Biostatistics Research.

References

[1] Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.*, **21**, 593–600. (First attempt to derive the asymptotic distribution of 2×2 contingency tables partitioned by empirical medians.)

[2] Borkowf, C. B., Gail, M. H., Carroll, R. J., and Gill, R. D. (1997). Analyzing bivariate continuous data grouped into categories defined by empirical quantiles of marginal distributions. *Biometrics*, **53**, 1054–1069. (Derived the asymptotic EBQP distribution. Source of the NS distribution.)

[3] Fleiss, J. L., Cohen, J., and Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psych. Bull.*, **72**, 323–327.

[4] Harkness, W. L. (1965). Properties of the extended hypergeometric distribution. *Ann. Math. Statist.*, **36**, 938–945.

(CONTINGENCY TABLES
EMPIRICAL BIVARIATE
QUANTILE-PARTITIONED
DISTRIBUTION
MEASURES OF AGREEMENT
QUANTILES
TWO-BY-TWO TABLES)

CRAIG B. BORKOWF
MITCHELL H. GAIL

NONLINEAR ESTIMATION, MAK'S ALGORITHM FOR

In many statistical applications, the parameter vector $\boldsymbol{\theta}$ of interest is estimated by a root $\hat{\boldsymbol{\theta}}$ of a possibly nonlinear equation $f(\mathbf{y}, \boldsymbol{\theta}) = 0$, where $\mathbf{y}$ is the vector of observed data (*see* NONLINEAR MODELS). A well-known problem of this kind is that of maximum likelihood estimation*. Unfortunately, the equation $f(\mathbf{y}, \boldsymbol{\theta}) = 0$ may have no explicit solution, so that iterative numerical methods are required. Some popular algorithms are Newton–Raphson*, quasi-Newton, and Fisher scoring (*see* SCORE STATISTICS). Based on the concept of conditional expectation, Mak [4] proposed a new approach, easy to implement, which has found a number of applications [3, 6, 5]. The following short exposition is mainly based on Mak [4] and Mak et al. [5].

ALGORITHM

The Newton–Raphson algorithm converges quickly, but an analytical expression for the vector of derivatives $\partial f/\partial \boldsymbol{\theta}$ may not be easy to obtain. A major problem with the quasi-Newton method in statistical applications lies in the numerical instability of the iteratively updated $\partial f/\partial \boldsymbol{\theta}$. In Mak's approach, a sequence of values $\{\boldsymbol{\theta}_{(r)},\ r = 0, 1, \ldots\}$ is constructed, which converges with a probability approaching 1 to $\hat{\boldsymbol{\theta}}$ from any starting value $\boldsymbol{\theta}_{(0)}$. But, unlike Newton-type methods, the construction of the $\boldsymbol{\theta}_{(r)}$ sequence does not involve $\partial f/\partial \boldsymbol{\theta}$.

Let $\mathbf{y}$ be an $n \times 1$ random vector of observations, and $p(\mathbf{y};\boldsymbol{\theta})$ its corresponding density function, where $\boldsymbol{\theta}$ is a vector parameter. The maximum likelihood estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is therefore obtained from solving $f(\mathbf{y}, \boldsymbol{\theta}) = 0$, where $f(\mathbf{y}, \boldsymbol{\theta}) = \partial \ln p(\mathbf{y}; \boldsymbol{\theta})/\partial \boldsymbol{\theta}$. Then:

(a) Fisher's information matrix is given by

$$\left.\frac{\partial g(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta})}{\partial \tilde{\boldsymbol{\theta}}}\right|_{\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}},$$

where $g(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathrm{E}_{\mathbf{y}}[f(\mathbf{y}, \boldsymbol{\theta}) \mid \tilde{\boldsymbol{\theta}}]$, and $\mathrm{E}_{\mathbf{y}}(\cdot \mid \tilde{\boldsymbol{\theta}})$ is the customary notation for the expectation taken under the density $p(\mathbf{y}; \tilde{\boldsymbol{\theta}})$ of $\mathbf{y}$.

(b) Suppose $\boldsymbol{\theta}_{(r)}$ has been given. Then we define in the $(r+1)$th iteration $\boldsymbol{\theta}_{(r+1)}$ as a root of the equation (in $\tilde{\boldsymbol{\theta}}$)

$$g(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}_{(r)}) = f(\mathbf{y}, \boldsymbol{\theta}_{(r)}); \qquad (1)$$

then

$$\boldsymbol{\theta}_{(r)} \to \hat{\boldsymbol{\theta}} \quad \text{as } r \to \infty.$$

Furthermore, $\boldsymbol{\theta}_{(r)} - \hat{\boldsymbol{\theta}}$ is $\mathrm{O}_p(n^{-r/2})$.

Result (b) implies that if the equation

$$g(\boldsymbol{\theta}_{(r+1)}, \boldsymbol{\theta}_{(r)}) = f(\mathbf{y}, \boldsymbol{\theta}_{(r)}) \qquad (2)$$

can be solved explicitly, the algorithm in (b) can be easily implemented and a high degree of accuracy is obtained in very few iterations. When (2) does not have an explicit solution, Mak [4] suggests the linearization

$$g(\boldsymbol{\theta}, \boldsymbol{\theta}) + \left(\left.\frac{\partial g(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta})}{\partial \tilde{\boldsymbol{\theta}}}\right|_{\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}}\right)'(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta})$$
$$= \left(\left.\frac{\partial g(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta})}{\partial \tilde{\boldsymbol{\theta}}}\right|_{\tilde{\boldsymbol{\theta}}=\boldsymbol{\theta}}\right)'(\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}) = f(\mathbf{y}, \boldsymbol{\theta}). \qquad (3)$$

Thus $\boldsymbol{\theta}_{(r+1)}$ is the solution to the linear equation (3) (with $\boldsymbol{\theta}$ replaced by $\boldsymbol{\theta}_{(r)}$).

EXAMPLE

This is the multinomial problem discussed in Mak [4] and considered by Dempster et al. [2] to introduce the EM algorithm*. The data consist of a vector of counts $\mathbf{y} = (y_1, y_2, y_3)$, observed to be $(38, 34, 125)$. It is postulated that

ENCYCLOPEDIA OF STATISTICAL SCIENCES

UPDATE VOLUME 3

Eds Kotz S, Read C B and Banks D L

A WILEY-INTERSCIENCE PUBLICATION
John Wiley & Sons, Inc. 1999
NEW YORK • CHICHESTER • WEINHEIM • BRISBANE • SINGAPORE • TORONTO